

2

Integrated Health Information Architectures – Operationalising the Vision

At the conclusion of Chapter 1, we understood the challenge to conceptualise the linkages between a social system perspective and the technical parts of the architecture, such as the 'data warehouse'. In other words, the preceding chapter left us with the question of how to conceptualise a socio-technical architecture. Well, the answer describes the IHIA, in three layers:

- The social system level or the user level, which includes users, business processes and information use.
- The application level, which includes the applications and systems.
- The data level, which includes processes of data interchange and interoperability.

The principle of layers (with each layer providing services to the layer above it), is beautifully demonstrated by the seven-layer Open System Interconnection (OSI) protocol model. The OSI model is a framework for implementing communication protocols for data exchange in layers, ranging from the physical layer of bits and cables to the application layer (7) communicating directly with the application. However, while the OSI model is a formal model of protocols for data exchange, our three-layered model is a conceptual model, which we place on top of the technical layer of cables and specific communication protocols. While the underlying technical layers, such as the seven-layer OSI model, are strictly concerned with the syntax of data flows, our lowest level, the data level, which builds on services from the technical layers, is where the semantics of the data is brought in.

Schematically, this three-level architecture is presented in Figure 2.1 and outlined in Table 2.1 and then further discussed.

2.1 Level 1 of the IHIA: Users and Information Needs – Social System Level

One of the first challenges arises when constructing a IHIA, which includes all subsystems, is identifying the key users and relative perspectives and vision. This requires a clear understanding of the work processes to be supported and the users' needs and requirements. We have therefore, labelled Level 1 as the social system comprising of the user and the institutional use of information. This characteristic is what makes this level undoubtedly, the defining level of the architecture.

To follow on all the users is meaningless, as it does not aid focus and vision. If you ask stakeholders, what the key components and focal points in a 'good' HIS are, you

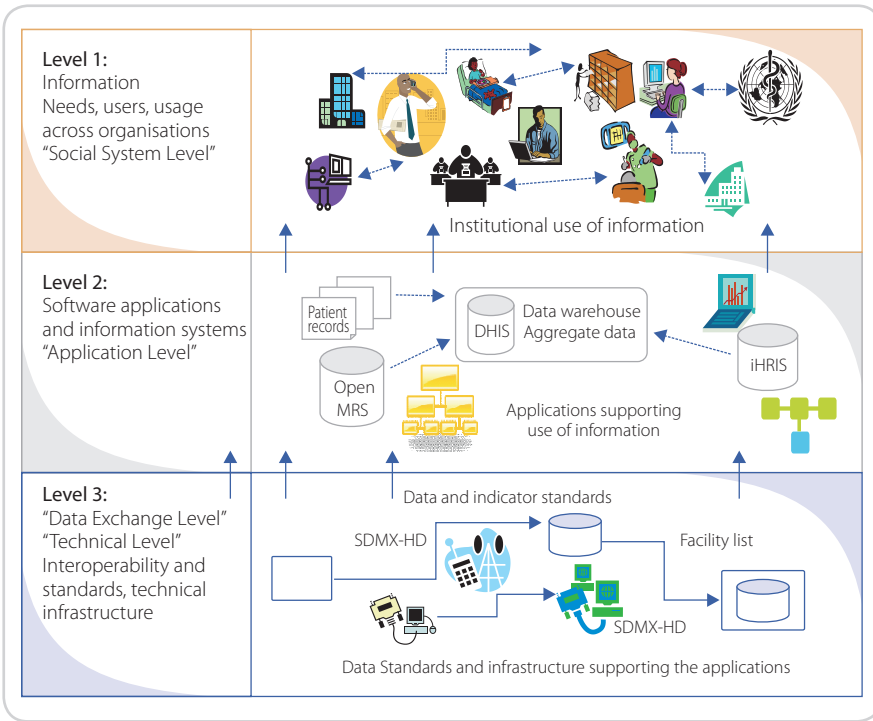


Figure 2.1 Three levels of enterprises architecture

Table 2.1 Three levels architecture	
Three Levels of the Health Information Architecture	
Level 1: Information needs, users and usage "Social System Level"	The users' information needs and actual usage of information; the business processes and functionalities to be supported by the HIS. Documented through users specifications and requirements within the context of the relevant business processes and organisational. The defining layer of the architecture!
Level 1 uses services from the level below (level 2)	
Level 2: Software applications and information systems "Application Level"	Applications and systems responding to the users' needs and providing the needed information and services to the users. Documented through SW application documentation, manuals, and actual implementations!
Level 2 uses services from the level below (level 3)	
Level 3: Data exchange, interoperability and standards "Data Level"	The technical level of data exchange and interoperability; the glue of it all, data and technical standards for interoperability of data between systems and applications, enabling data flow. Types of standards described differently, from formal standards for data exchange to data dictionaries of data standards and semantics.

will get different answers depending on the respondents' roles and 'locations'. A health technician in the hospital would emphasise medical records, a manager may emphasise information needed for decision-making, and an epidemiologist would point to population-based surveys, and so forth. Typically, public health practitioners and health managers tend to focus on HMIS type of systems; indicators and aggregate data supporting management and decision-making, whereas clinicians and medical practitioners, and IT technicians will tend to focus on electronic medical records. As argued earlier, the design focus of the IHIA is on information for management, decision-making to support effective health services delivery at all levels of the health system, related sectors, and institutions, as well as information dissemination to the general public.

Need for information, for decision-making and management may be described by the availability of essential information from across the sectors in appropriate formats, so that comparisons, trends, and correlation between areas, across sectors and over time can be carried out. Different levels of the health system will have varying needs, according to the role they play in the system and their scope or power of decision-making. The following types and areas of user needs are important:

- The need for integrated information, so that the information that originates from different areas is available at 'one point'.
- Different levels and types of management in the health sector have varying needs. For example, the lower levels need richer and more granular data, while the higher levels need less data in a more aggregated form.
- Information for action, where the focus needs to be on essential data and indicators, which are linked to targets and real usage.
- The form and presentation of the information, and how to access it. For example, reports, graphs, maps, statistics – real time and online using different media.

Each of the above principles are now elaborated.

The Need for Integrated Information

Fragmentation is a major problem faced by HIS in most developing countries. Countries are typically saddled with historically institutionalised 'vertical' programmes, each collecting 'their individual' information independent of other programmes, with little regard to supporting the overall HIS. The result of this is excessive data, with great overlaps of the same data collected many times, and sometimes leaving large gaps in critical areas with no data at all. Furthermore, data and indicators are poorly standardised and represented in incompatible formats, making comparisons and analysis across programme areas, difficult. Some programmes such as HIV/AIDS are 'strong' and collect a lot of quality routine data, supported by regular surveys, whereas, other programmes such as environmental health are 'weak' and collect little or poor quality data. Here is an example from Zanzibar (Box 2.1).

The need for integration of information systems and to have information from different sources available at 'one point', as expressed here, are requirements that are expressed by many countries – a need to move from 'fragmentation to (more) order'. This is the aim of various national health reform efforts, advancing in various countries such as in India – to have information from different areas 'at one point' and at 'the press of a button'.

Box 2.1

An example of fragmentation in HIS from Zanzibar

Fragmented HIS: An example from Zanzibar

The situation in Zanzibar was as described above, characterised by fragmented and uncoordinated HIS. At a startup meeting of a project in 2006 to integrate and strengthen the national HIS, the Director of Health said:

“When I need to get an overview of the situation across different diseases and services areas, nowhere is that information available. I have to ask for information from a large number of programme offices – Malaria, HIV/AIDS, EPI, hospitals, and so forth. The resulting information is not easy to comprehend, compile and analyse, as each office tend to structure their information differently, and it is difficult to get exactly the information I need.

What I need is to have all the important data from all offices available at one point, in one office, so that I can get it here on my desk – on my computer. By the press of a button.”

– Ministry of Health, Zanzibar, 2006

Varying Information Needs of Users and Administrative Levels

Data and information should support different work and management processes, at various levels of the health system; from patient management to the management of health facilities, and from districts, states and national levels. The management of patients and individual clients requires data on individual encounters with the health services; on diagnosis, procedures, laboratory tests, results, and so forth. The amount of data collected on each individual is significant and has implications both on the database and server capacity. Similarly, in a community setting, huge amounts of data are required to be collected to enable the registering and tracking of services to all pregnant women or children immunisation in a district, or in a hospital. In order to manage a health facility, aggregate data is required, for example, data such as the number of pregnant women taking part in the antenatal clinics, and the percentage of them that are not following up on their check-up schedule. In order to rectify a problem identified by a low percentage of women not coming for their check-ups, the level below – the outreach worker – would require a list of names and addresses of the pregnant women to be visited by health workers, at the health facility level. More generally, we may say that while patient management requires individual ‘name-based’ record systems, health facility management would need aggregate data and indicators on how the various programmes and activities are performed with respect to their targets. Most of this data can be aggregated from register books and individual records, all of which are either paper or computer-based.

If we move up from the outreach worker and the health facility, to the district, we see that the data and indicators needed for managing will be even less granular, requiring more limited amounts of data. The reason is that, while facility management is directly engaged in delivering services to clients and patients, the districts are located ‘one step up’ and require managing basically through the facilities in the district. The district, therefore, requires data on how each facility and health programme (not individual

patient) is performing, which will be the aggregate of the data generated at the facility. Moving from the district to higher levels of the state and national level, even lesser data of more minimal granularity is required. This principle of the data needs in all the levels and across, is presented in Figure 2.2:

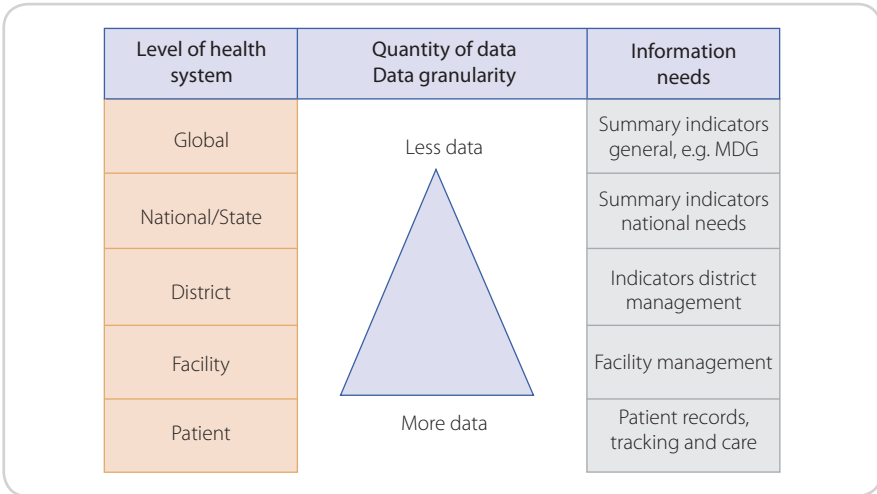


Figure 2.2 Data needs in all the levels and across

In line with the Figure 2.2, Table 2.2 outlines management functions and the corresponding information needs by the various levels in the health system.

Management functions	Information needs
Facility management in district health centre services delivery: <ul style="list-style-type: none"> • First level curative care. • Antenatal care. • Deliveries. • Immunisation. • Under-5 care. • Chronic patient care and follow-up. • Outreach activities. • Community participation. • Drug and vaccines management. 	Health centre HIS: <ul style="list-style-type: none"> • Aggregate data summary of all activities in the facility, by month; number of clients /patients by category. • Indicators linked to targets for immunisation, etc. • Community profile and target population information. • Patient /client records (registers) on all categories. • Tracking (as a minimum) of immunisation, pregnant women – deliveries, chronic patients.
Hospital services and management: <ul style="list-style-type: none"> • Curative (referral) care. • Outpatient. • Inpatient. • Surgery. • Laboratory. • X-ray. • Drugs. • Resources management; finances, staff, buildings and equipment. 	District hospital HIS: <ul style="list-style-type: none"> • Aggregate data summary by month from all wards, specialties and services. • Indicators on death rates, average length of stay, bed occupancy, etc. • Financial data; consumption/income against budget. • Drugs; stock vs consumption. • Patient records system for quality and continuity of care.

District level management: <ul style="list-style-type: none"> • Manage health programmes; immunisation, antenatal care, nutritional activities, under-5 care, family planning, HIV/AIDS, etc. • Manage districts resources; human resources, finance, equipment/ buildings, etc. • Supervise health facilities in district and provide feedback. 	District HIS: <ul style="list-style-type: none"> • Aggregate data summary on all activities and health programmes by month by facility. • Monitoring all key indicators. • Community profile and target populations – denominators. • District maps on demographics and health services. • Employee records and human resource summary data.
<ul style="list-style-type: none"> • Develop/evaluate annual work plan. • Organise health promotion in community and other activities. • Manage ambulance services, vehicles, transport and logistics. 	<ul style="list-style-type: none"> • Financial data – summary; consumption against budgets.
State (province, region) level: <ul style="list-style-type: none"> • Co-ordinate and plan state and regional health services and programmes. • Manage resource allocation. • Supervision of districts. • Monitoring and evaluation of district and programme performances. • Logistical support. • Human capacity development – planning. 	State HIS: <ul style="list-style-type: none"> • Monitoring all key indicators from health programmes and services in state. • Monitoring of budget utilisation with respect to programme performance. • Analysing inter district performance to identify which districts and programmes need more or less support. • Employee records and human resource summary data. • Financial data – summary; consumption against budgets.
National level: <ul style="list-style-type: none"> • Policy development. • National planning. • Evaluation and impact analysis of various programme evaluation. • Resource allocation. • Monitoring and evaluation – all programmes, services and health, demographic, and socioeconomic status. • Human capacity development – planning. • Epidemiological analysis. • Research agenda definition. 	National HIS: <ul style="list-style-type: none"> • Monitoring all key indicators from health programmes and services across states. • Monitoring of budget utilisation with respect to programme performance. • Analysing inter state performance to identify which states and programmes need more or less support. • Employee records and human resource summary data. • Financial data – summary; consumption against budgets for states.

In the context of various information needs by different levels in the health system, some important distinctions need to be understood.

Individual level data, from health programme registers (for example, Mother and Child register book) are represented either as individual record systems on paper, as rows in primary registers, or computer-based medical record systems. Clients registered in the register book during a month will typically be compiled at the end of the month to provide a total number of antenatal clients (ANC) as raw data.

Raw data are aggregations from individual data, to make sense in a particular context, representing the absolute achievements in a facility over a determined period. For example, a raw data form provides the data that 75 infants were given BCG vaccine, and 53 were given measles vaccine, in a particular PHC, in a month. From the raw data, it is difficult to make an inference of whether or not this is a good or poor achievement. Thus, it is vital to see how the data relates to the

target population; in this case, how many total children were expected to be given BCG or Measles vaccines.

Indicators are quantitative measurements to indicate/show particularly significant events or conditions, and to measure its changes, directly or indirectly. Indicators are typically based on processed data elements, and will generally include a numerator and a denominator. The numerator is a count of what is being measured and the denominator will typically be the size of the target population related to the count. Indicators are generated by calculating data according to a formula. While the rate of immunised children in a district represents a calculated indicator, the count of cases of maternal death may serve as a non-calculated indicator.

Information needed for management, decision-making and, monitoring and evaluation, will typically be indicator based, of which at least four types can be identified (Box 2.2).

Box 2.2

Four types of indicators¹

Types of Indicators

Count indicators: Measures the number of events without a denominator.

Example of count indicator = Number of recorded maternal deaths

Proportion indicators: Typically expressed as a percentage. The numerator is part of the denominator.

Example of proportion =
$$\frac{\text{Number of first ANC visits within first trimester}}{\text{Total number of ANC visits including within first trimester}}$$
 indicator

Rate indicators: Measures the frequency of an event, during a specified time, in a specified population, usually expressed per 1,000, 10,000 or 100,000 population. Used to measure the probability or risk of, for example, infant mortality (typically per 1,000) or disease in a defined population.

Example of rate indicator =
$$\frac{\text{Number of infant deaths}}{\text{Total number of live births}} \times 100$$

Ratio indicators: The numerator is not a part of the denominator.

Example of ratio indicator =
$$\frac{\text{Number of female HIV positive}}{\text{Number of male HIV positive}} \times 100$$

¹ After Lippeveld, Sauerborn, Bodard, World Health Organization, 2000

2.1.1 Information for Decision-making and Focus on Essential Data

It was an academic trend in the early nineties that advocated decision-making systems to be rigid, especially on what information it includes in the scope of decision-making; for example, if you have no influence on a situation, you should not include data from that area. This principle translated into the influential Minimum Datasets (MDS) approach, developed in the post-apartheid era of South Africa. In the MDS approach, each level should define their most important data and indicators to satisfy what the levels below would need, to collect the specified data. When applying this approach

in other countries, however, there seemed a recurring problem of creating the MDS, because of the strong central mandate. This also led to the problem of the data being reported through the official HMIS to be narrow as it catered primarily to the needs of the national statistics division. Thus, health programmes had to initiate their own additional data reporting systems, which created overlapping and confusing structures of reporting. The more general strategy emanating from these experiences was to focus on involving various stakeholders in defining information needs, and within a more integrated framework.

Information needs for management and monitoring, are quite similar in type and areas of data, but data typically used for ‘M & E’ – monitoring and evaluation – often goes beyond the normal scope of decision-making for the health services. Here more comprehensive data, for example, on disease patterns, socio-economics, infant mortality, vulnerable groups and equity would be analysed and used despite not being directly applicable within the health services or the scope of power. However, if advocated correctly, such more comprehensive information could translate into political action. In this way, information from outside the scope of decision-making power may be turned into use. As a result, excess or overflow of data, poses a problem. This represents the design problem of creating a balance between a strictly action-led approach and a more comprehensive approach, allowing for richer and more overall information needs. An attempt to capture this dilemma, can be seen in Figure 2.3.

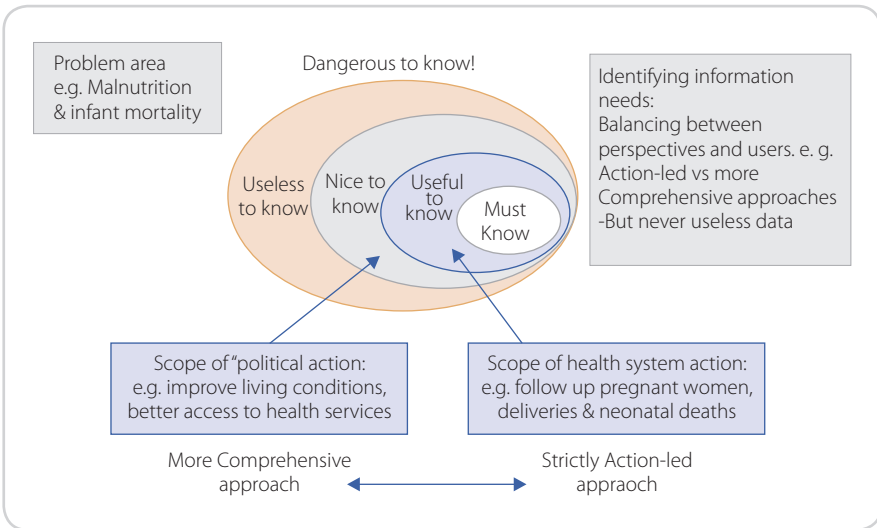


Figure 2.3 Balancing varying information needs

The HIS framework described above, depicts the different kinds of information in terms of its utility. For example, the essential information (‘must know’) for follow-up on health action, and the ‘useful to know’ information, that is required for public dissemination and advocacy. At the time of overflow of data and information, essential data and indicators are taken as information to address the needs of the user groups. For example, are indicators measuring achievements against relevant targets, rather than indiscriminate amounts of ‘raw’ data. This issue is further elaborated.

2.1.2 Information – For Purpose, For Action!

It is important to understand that all organisations, public and private, in both developed and developing countries collect more data than they actually use, for decision-making and action. Some data is used just as a control, to ensure the job is done, while some data collection is simply done to satisfy bureaucratic inertia. A review done in South Africa showed that more than 50% of the data submitted was neither looked at nor captured, but dumped. In most countries, where HIS are less mature, the percentage would be considerably higher.

The countries assessed as part of a HMN project to develop an assessment tool in 2005, showed similarly that a significant amount of data being collected was not analysed or used. Yet, there were encouraging aspects, particularly in South Africa and Thailand:

- In South Africa, individual provinces, districts and national programmes were systematically using information for action, with proven successful outcomes. For example, the health sector was in a major shift towards activity-based budgeting where information from the HMIS would be a deciding factor on resource allocation.
- In Thailand, there was a maturity around the use of information, which could be gauged from the fact that the information was being used to support the everyday working of the health department, as well as, national initiatives like 'Healthy Thailand', and the national insurance scheme.

Both countries, despite significant and increasing use of information for action, still collected more data than they used, and health staff continued to face major challenges in increasing analysis and use, especially at the district and local levels.

The HMN study reported that in most other countries, such as India and Vietnam, the relation between data collection and information use for action was severely mismatched.

- For example, in the Indian states of Andhra Pradesh and Karnataka, each Primary Health Centre (PHC), collected more than 1,000 data items on a monthly basis, including many duplicates and data from programmes that had been discontinued. Indicators were hardly used, and targets were uniform based on population figures rather than on programmes for action, which took account of inequities in population, social status and geography. The targets, mostly unrealistic and unreachable, were a constant source of manipulation as the health staff tended to 'doctor' figures to show that they had achieved the targets because the punishment for not reaching targets was way worse than the punishment for being caught lying. The strong culture of reprimand, which existed in the health sector, served as a strong deterrent for staff from reporting performance that may invite reprimand. Further, the numbers reported tended to be 'perfect', showing achievement of targets, indicating a non-believable character to the numbers.
- Vietnam too, had a centrally planned health service, and data was commonly manipulated to artificially achieve impossible targets, without any cross-checking from supervisors or top management.

One prevailing and important problem in most countries is the vicious cycle; of poor quality data not being used; and because it is not used, remains poor in quality. The poor quality is perpetuated by weak mechanisms for supervision, as well as the

multiple and outdated population figures that are used. Many clinics deal with multiple sets of population figures – of the health department, the revenue department, and figures compiled by household surveys conducted by independent agencies.

The ‘holy grail’ of promoting analysis and use of information, is clearly to link information to resource allocation (budgets), as well as develop indicator-driven short and medium-term planning. Linking HIS indicators to targets is in itself no guarantee for the collection of quality data or for its use in decision-making. Fulfilling targets can become a meaningless propaganda exercise, where everybody is making up data, but nobody wants to ‘rock the boat’. Concrete planning and targets are crucial, but plans and targets must be realistic and linked to resource allocation. Evaluation of the achievement of targets should focus on the entire process of the HIS, and not just the outputs. Evaluation should stimulate constructive suggestions on how to improve the process, rather than on providing reprimand and public shame.

2.2 Level 2 of the IHIA: Software Application and Information Systems – ‘Application Level’

The information requirements for the HIS, for different administrative levels of the health services were outlined in the previous section. Key management needs that were identified – to have access to the essential data and indicators, integrated from across areas at one point, in one system to enable analysis, display and presentation of data using maps, graphs and tables. Further, it should be easy to design and customise reports locally, and more generally, the system should be easy to use. These requirements as we have argued are best addressed through a ‘data warehouse’ approach, whose software application level is now discussed.

A ‘data warehouse’ comprises of flows of data from various health services and programmes, health facilities, school health, and community. These flows are conceptually led through and managed by the district office and their part of the data warehouse, where data needs to be analysed; and feedback to those who have reported it, with a view to strengthen the management structures in the district and levels below. Further, there is an upward flow of reports that are required by the levels above. In the HMN description of the ‘data warehouse’, data of different types are imported from other computer-based systems such as, census, medical records, financial systems and so forth. In our context, however, in many settings the data collection, management and reporting at the local level will still be predominantly paper based. Data capturing from paper forms directly into the ‘data warehouse’, or the HMIS, is currently the typical situation. However, we are in the beginning of a process of substantial changes in this regard, where the primary management of data will also be computer-based. In other cases, we also see the use of mobile phones to send data into the warehouse through a SMS. While we see these winds of change in certain contexts, what we need to plan for, is a combination of environments, wherein paper, computer, and also other devices like mobiles where they are used. An example is provided from the state of Himachal Pradesh, India, reflecting an evolutionary approach to ‘data warehousing’ and HIS development (Box 2.3).

Here, paper-based systems are gradually being computerised and new types of data are being included in the ‘data warehouse’, with the point of departure being the current HMIS.

Box 2.3

Designing and implementing the data warehouse for Himachal Pradesh, India

The Data Warehousing Approach – Example from Himachal Pradesh, India

Himachal Pradesh is a state in Northern India, which is one of the more progressive states in the country with respect to its basic health indicators, for example, related to maternal and child health. Since the routine systems are relatively well-stabilised and institutionalised, the state has recently embarked upon a more sophisticated initiative of creating an IHIA with technical development being carried out through a national NGO called Health Information Systems Programmes (HISP), India.

This IHIA is comprised of a set of distinct information systems development processes, which are all integrated through the overall 'data warehouse' approach and 'glued' together through the application of standards. For example, the new WHO initiated, shared standard for data exchange and interoperability called SDMX-HD:

- *Building district and state data warehouse, based on the current routine HMIS:* The basic part of this data warehouse, is already in place with the HMIS database (based on DHIS2) for routine data, where the state carries out data entry at the sub-district level. The initiative in process, is to include more types of data and to extend the data structure of this 'data warehouse' by including data on the two lowest level facilities – the PHC and the Sub Centres reporting to this PHC – where till date this data was being aggregated and entered at the sub-district level of the block, which is a group of PHCs. Further, this system forms the framework within which the other projects, described in the following, are being integrated by feeding aggregate data to the 'data warehouse' and using data from it.
- *Mobile telephone reporting from Sub Centre and community levels:* Currently, the Sub Centres are reporting their data on a monthly basis, on a number of paper forms to the above HMIS system. Given the harsh winter, isolating many valleys in the state during the winters, the project is to use a mobile phone to report data directly and seamlessly to the 'data warehouse'. This mobile based Sub Centre reporting has been previously piloted in one block of the state, with positive results. Taking this many steps further, the mobile application is being extended both geographically and functionally. The geographical scaling plan involves extending from one block in one district to the complete district, and then to the other 12 districts in the state. The functional scaling is being done, by including all the formats that the field nurse [called the Auxiliary Nurse and Midwife (ANM)] has to report in a month (nearly 10 different formats) on to a mobile phone, in contrast to the earlier application, which only includes one dataset.
- *Name-based tracking of pregnant women, deliveries and children for immunisation – the DHIS Tracker system:* The government of India, has recently initiated a programme of name-based tracking, where all pregnant women would be tracked by names over the lifecycle of the services of antenatal,

delivery and postnatal care, and also all newborns over the lifecycle of their immunisation. The software to support this tracking (called DHIS Tracker) has been developed by the HISP network as a module in the DHIS2. The system is planned to be implemented in the state in a phased manner. From this system, the routine monthly reports on number of ANC visits, deliveries and immunisation will be aggregated, transformed into an adopted standard, and exported to the 'data warehouse.' Over time, the DHIS Tracker data would also be transmitted through a mobile phone application.

- Comprehensive hospital information system – integrating electronic medical record system and 'data warehouse' for hospital management: This project is two-folds.
- To develop an electronic medical record system for district hospitals, one which is easy to adapt to the needs and available resources, even in smaller hospitals. The system needs is flexible and has been scaled from initial admission and billing modules to a total of 10 modules and functionalities, as learning and human resources have also developed.
- Integrate the aggregate data from the medical record system with other types of data from the hospitals, needed for management, such as human and other resources, finances, laboratories, and drugs, in what we are labelling a 'data warehouse' of aggregate data for hospital management. Further, to extend this system, to include all patient related data, corresponding to services provided to individuals, in the setting up of 20 district hospitals in the state.

There are particular informational needs of hospital and facility management, which are not covered by medical records systems, since they primarily target patient management. The 'data warehouse' for hospital management can be understood, as in the district, by conceptualising wards and specialties as facilities in a district; organising aggregate data by wards and specialties; and correlating the data with the number of beds, staff and other services, thereby being able to analyse and present key hospital indicators, such as:

Bed occupancy: Number of patient days/nights divided by number of beds, typically provided by month; bed nights during a month divided by number of beds \times 30.

Average length of stay: Number of patient nights divided by number of discharges, typically by month.

Death rate: Number of deaths divided by number of patients; by age, service and ward.

Infection rate: Hospital infections divided by number of patients; by ward, age and service.

- *Geographical Information System (GIS) – presentation of data using maps and other data representation tools:* The geographical co-ordinates on each health facility, and the borders of Districts and Sub Districts, have been included in the 'data

warehouse’, and the GIS module in the DHIS2 will enable the mapping of services, health and demographic status related to facilities, districts and other geographical boundaries.

The data presentation modules of the ‘data warehouses’, will more generally be developed further, to include a dashboard for easy access to graphical tools (bar charts), mapping, as described above and tabular and pivot enabled formats.

The next phase of the project will cover the creation of a *Human resource management information system*, which will include records of the employees in the health services. Aggregate data on number of staff and qualification by health facility, represents important information for managing hospitals, districts and more generally the health services in the state, and will be imported into the ‘data warehouse’. Human resource data by facility also represents important data for the mapping functionalities in the GIS.

The IHIA – Integration of the Projects and HIS

All the above initiatives, distinct projects and sub-systems of the HIS, are being integrated within the framework of the integrated ‘district-based’ state data warehouse, which may be conceptualised as a development and extension of the former HMIS, where the aggregate data from all hierarchical levels of the health administration come together. Data standards for exchange of aggregate data will be used to feed data from the various systems to the ‘HIS data warehouse’, which is being created at all levels of the administration; state, district – and to support management in the hospitals. The scheme for integration includes the following:

- The data reported by mobile telephones from the Sub Centres are being directly fed into the ‘data warehouse’, and will gradually make the reporting on paper and following data capture, obsolete.
- The name-based data tracking will, first, be dominantly reported on paper and captured in the DHIS Tracker database at block level; and second, gradually be registered and sent by mobile phones to the DHIS Tracker application. Aggregate data on pregnancies and ANC services, deliveries and immunisations will be fed into the data warehouse. The corresponding aggregate data currently being reported by paper and entered directly into the ‘data warehouse’ would be eliminated over time.
- Data from the hospital information system at the patient level, will be aggregated and exported directly into the ‘data warehouse’.
- All data and indicators will be available for presentation and analysis through the GIS and other reporting tools.

In this way, integrated data from different projects are being made available to a common set of reporting tools, for generating all required indicators to support management analysis and reporting at different levels of state, district, sub-district and hospital.

In the example of Himachal Pradesh, the 'data warehouse' represents an integrated framework – an 'umbrella', within which, the various systems and also data warehouses are gradually being plugged in and subsequently scaled. The existing HMIS and routine paper-based reporting, forms the backbone and point of departure. Here is a step-by-step process involved:

First Step: The data structure in the state data warehouse, previously called the HMIS database, is extended to include Sub Centres, which is important to strengthen data quality, data analysis and information use.

Second Step: It enables data reporting from Sub Centres, with the use of mobile phones. This is especially useful in the state, where parts of it are closed during winter as snow makes physical travel for data reporting problematic. The mobile network covers most of the state with all the Sub Centres having access, yet some may need to go to the nearest PHC, in order to submit their reports by SMS, or using the GPRS network.

Third Step: The electronic medical records in the district hospitals starts with patient admissions aggregate to developing management reports, such as for discharges and billing. Monthly hospital summary reports will be aggregated automatically by the electronic patient record system and transferred to the data warehouse.

Fourth Step: Registration of each pregnant woman and her follow-up until delivery, as well as, the registration of every individual infant and the doses of vaccines given to them, is a large undertaking as, initially, paper forms will be filled out in each Sub Centre and then submitted to the block, where the data is captured in the database. Later, mobile phones will be used for reporting data, a system currently being developed in one district. Aggregate reports from this system will automatically be generated every month, or at different intervals, and sent to the data warehouse in co-ordination with other Sub Centre reporting. Furthermore, reports and schedules of pregnant women for deliveries will be communicated to the relevant hospital, and data will be transferred to the hospital database.

Fifth Step: A human resource management information system will be developed within the same framework, to feed aggregate data on human resources into the data warehouse.

Sixth Step: GIS, analytical and reporting tools are encompassing data from all the different systems through the 'data warehouse'.

The planning and development of the IHIA in Himachal Pradesh, illustrates the benefits of a shared architecture and integrated framework, represented by the 'data warehouse'. Without that, the different components would have been easily developed as totally separate/independent entities, with little or no interconnectivity. For example, two separate mobile projects, one hospital project, one GIS project, one HMIS project and one human resource management project, which would have contributed to further fragmentation rather than to strengthen integration leading to more effective monitoring and decision-making, as is the case here.

In Figure 2.4, the 'data warehouse' design from the above example of Himachal Pradesh is schematically illustrated as the application level of the architecture. The 'data warehouse' scheme in this figure, replicates for each administrative level of health services, from the larger health facility – the hospital, and upwards to the district, state and national levels. Obviously, the data entry based on paper will be at the local level, as well as, the data imported from the medical records and human resource system.

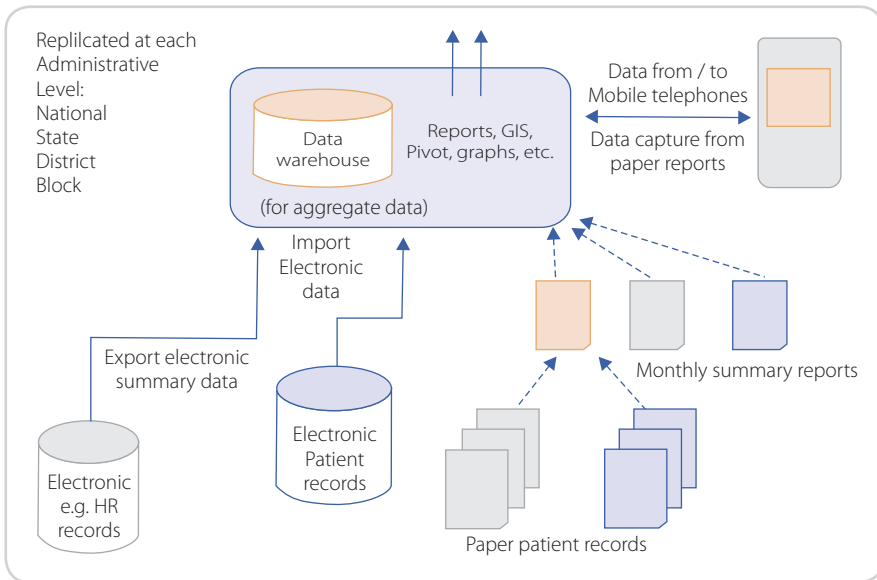


Figure 2.4 Conceptualising the Himachal Pradesh IHIA, and making it general

Generally, as pointed out earlier, the more local levels will need more granular data than the higher levels. Using a hospital as example, we see that a lot more data will be needed to manage the concrete local hospital, than needed at the state level; where 20 hospitals are being monitored.

2.2.1 Presentation and Dissemination of Information – The Purpose of the Data Warehouse!

Although discussions around data warehouses tend to be dominated by technical issues of data sources, and how to get the data into the warehouse, it is important to remember that the main purpose of this approach is to support the use of information; the output and Business Intelligence part of the data warehouse concept. This requires strong functionalities in the presentation and dissemination of information. Good presentation greatly enhances the use of information. While indicators as figures in a drawer are not very useful, displayed as graphs on the wall and updated every month, can make a difference. Some examples are now provided, of how the data warehouse should support different ways to present information – graphs, maps, charts and tables.

Example 1: This is the executive dashboard which provides the ‘birds’ eye view’ supported by the DHIS2 (Fig. 2.5). Using this functionality, the user can configure some key charts that he/she should see every time they log into DHIS2. For example, the user below has charts on immunisation, Malaria and ANC vs BCG coverage. Indicators of interest, thus, pops up for the user, giving a birds’ eye view every time on login.

Example 2: The Figure 2.6 shows how the GIS can be used to map indicators, in this case by Chiefdoms in Sierra Leone. The legend set, can be customised by the user, to define the ranges of display. The GIS package used here is the OpenHealthMapper –

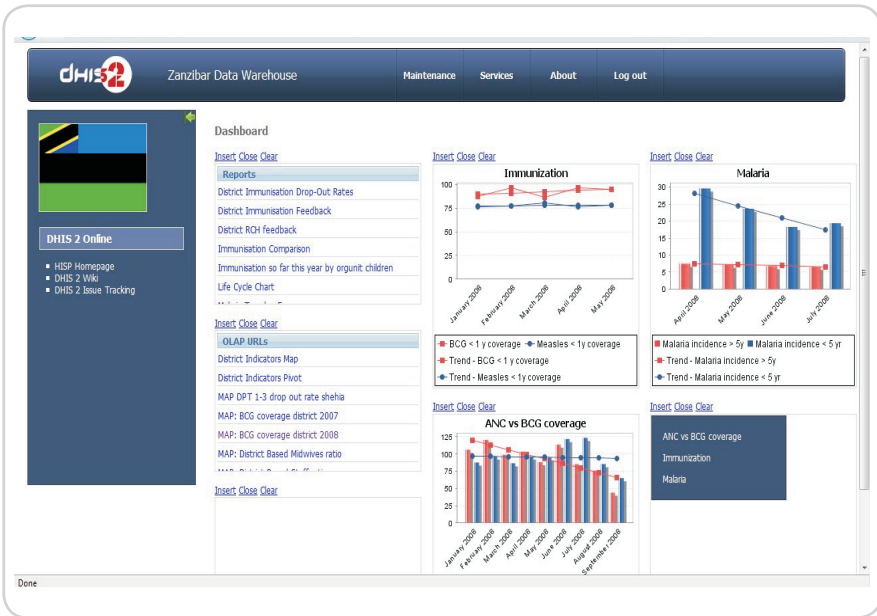


Figure 2.5 Bird’s eye view through the DHIS2

the web-based version of the earlier version of WHO’s Health Mapper. This GIS module is bundled in the DHIS2, which allows integration of the non-spatial data collected in the DHIS2 with the GIS shape files stored in the same database.

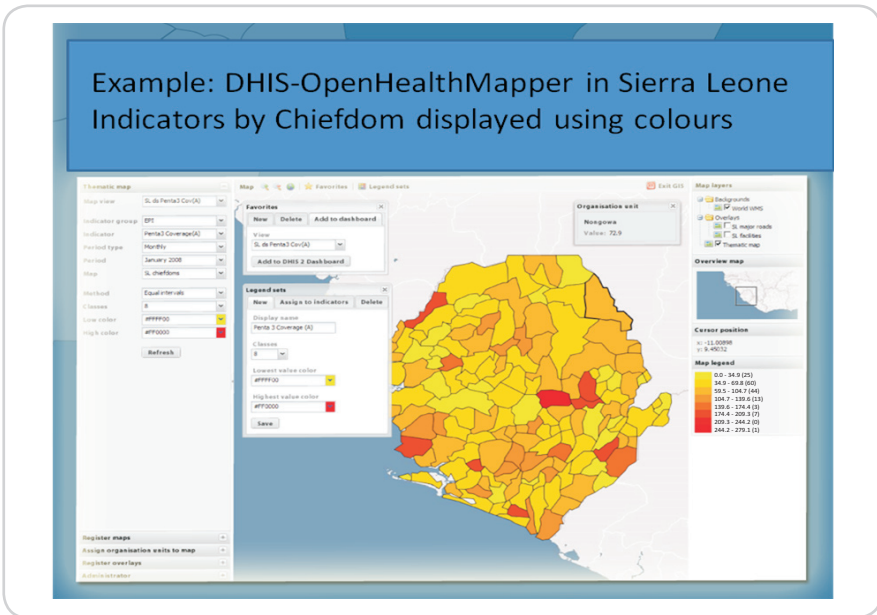


Figure 2.6 GIS presented indicators by Chiefdom in Sierra Leone

Examples 3 and 4: Presents two kinds of chart (Figs 2.7 & 2.8) representations, both taken from the DHIS2 deployment in India. The *first* chart compares ANCs first and third visit coverage, in the same graph for the different districts in the state. Seen together, these two graphs help to get an overview of the quality of ANC service care. The *second* graph compares home and institutional deliveries across districts in a state. Since, the sum of the two (home and institutional) must, by definition be 100%, seeing these figures in the same graph, helps to identify which districts; report high or low – home and institutional deliveries; and areas where intervention is required.

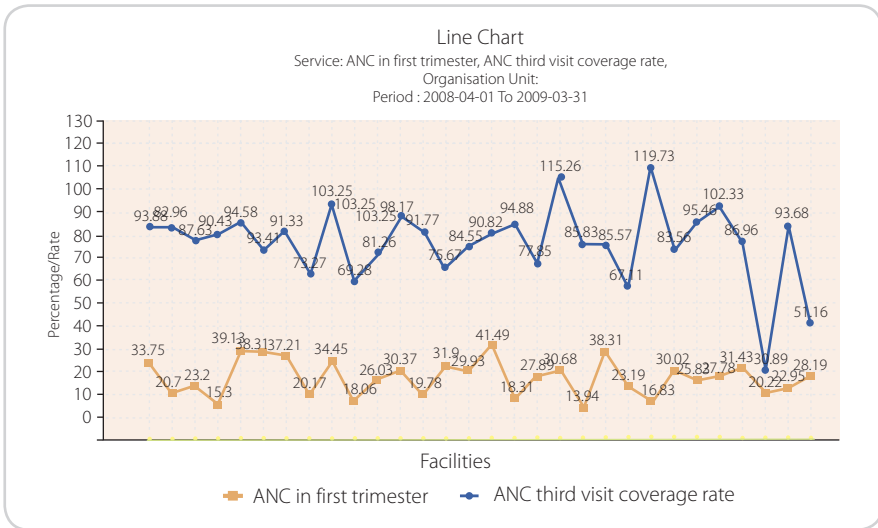


Figure 2.7 Line chart comparing ANC first trimester and third visit coverage

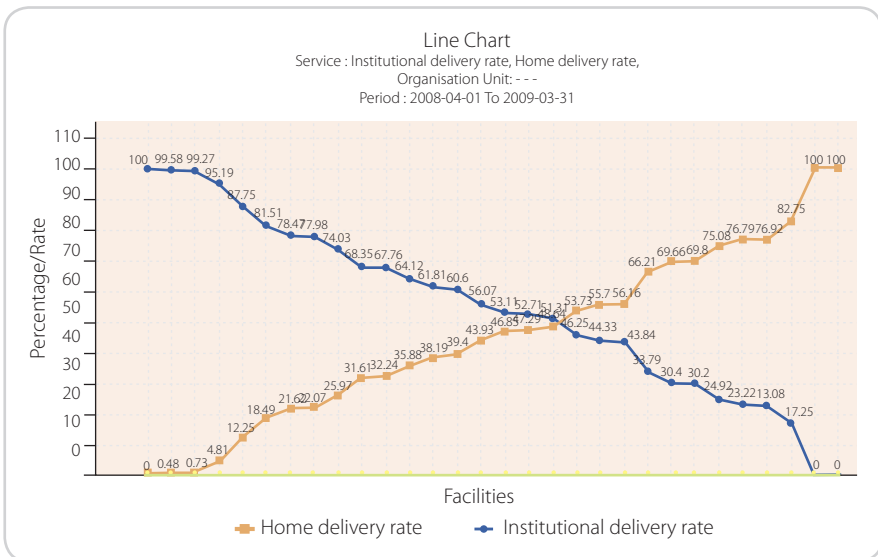


Figure 2.8 Line chart comparing home and institutional deliveries

Note: Names of the organisation units are kept anonymous in the above charts.

2.3 Level 3 of IHIA: Data Exchange, Interoperability and Standards – ‘Data Level’

While the Application level of the IHIA (level 2), may be depicted as consisting of a number of boxes, symbols and interconnecting lines, the data interchange (level 3) of the architecture, is mainly concerned with the lines connecting the applications, representing interoperability between the applications. In our context, interoperability refers to the ability of a system to use information or functionality of another system by adhering to common standards. According to the European interoperability framework (after Vernadat, 2005), interoperability can happen at three levels:

- *Technical level:* The ability of the ‘data warehouse’ to receive data from, for example, the medical record and human resource management system, and to exchange messages with these systems specifying which data to transfer. This is typically handled by technical protocols for data transfer. The effort of WHO to develop the SDMX-HD standard for data interchange is an important example in the contemporary public health context.
- *Semantic level:* This is linked to the definition, meaning and selection of the data to transfer between the systems. Data and indicator sets and definitions, data dictionaries, and the ICD 10 (International Classification of Diseases v10), illustrates semantic type standards used in our context.
- *Organisational level:* This is typically linked to the interaction of people and organisations, in order to define and decide which data and information should be shared and transferred, and the actual use of the data across organisational borders – interoperability between organisational units such as district and state hospitals.

While the technical and semantic levels of interoperability are both part of the third level, the data interchange level of the architecture model, interoperability at the organisational level is part of the first level, the level of use and users – the social system level. As interoperability refers to the ability to apply standards in a ‘practical way’, the three levels of interoperability are replicated in the world of standards. In the next chapter, we discuss notions of standards and standardisation, along these three levels in greater detail. Here, we focus on the two levels labelled technical and semantic interoperability, which together make up the third level of our architecture. In order for the different software applications to communicate and interoperate, there is a need for agreed standards; for what data to exchange; the format for exporting the data from one system; and importing the same data in another. This approach is further discussed.

2.3.1 Data Standards

Standards for health data, and other related data, will include both the definition of the data element; its content and meaning; and the name or reference. Table 2.3 is taken from the data and indicator dictionaries used in India, including the definition and rationale of two types of maternal health indicators, alongside the data elements needed to calculate them. To recall, data elements denote the disaggregated basic data elements, also called variables, counts and even indicators in other contexts, while indicators in most cases are calculations based on data elements alone or together, with demographic data, where the data elements are the numerators. Some examples of indicators are provided in Table 2.3.

Table 2.3 Example of Indicators

MATERNAL HEALTH INDICATORS – INDIA	
Indicator MH 1: Antenatal care first visit coverage rate A: ANC – First Visit B: ANC First Visit in first trimester C: ANC registered under Janani Suraksha Yojana (JSY)	
Definition	<ul style="list-style-type: none"> Percentage of pregnant women who used Antenatal Care (ANC) provided by skilled health personnel, for reasons related to pregnancy, registered in first trimester of pregnancy. This indicator is also known as 'Any antenatal care visit'.
Numerator	A: New registered/first ANC visit of a pregnant woman B: Pregnant women registered within first trimester C: New women registered under JSY
Denominator	A: Total expected pregnancies B,C: Total number of ANC registered
Rationale	<ul style="list-style-type: none"> This first visit should be a 'registration' visit, with all the initial procedures, relating to assessing/preparing a woman for pregnancy and delivery. This should include history, examination, initial blood tests and immunisation. ANC care coverage is an indicator of access and use of health care, during pregnancy. All women should have at least three antenatal visits during a pregnancy. ANC should start as early in pregnancy as possible. ANC registration percentage in the first trimester shows early care and level of awareness. Percentage of women registered under JSY shows: Number of women entitled to benefits under JSY. This includes: (a) All women in Empowered Action Group (EAG) and Northeast (NE) states. (b) Only Below Poverty Line (BPL) and Scheduled Castes and Scheduled Tribes (SC/ST) women. Percentage of pregnant women receiving any ANC is a sensitive indicator of outreach.
Data source	<ul style="list-style-type: none"> Antenatal/pregnancy registers; maternal health cards. Household surveys. Population data – an estimate of the number of pregnant women is close to the number of children born (2.2 to 3.2% of population).
Suggested level of use	National, state, district/block and Sub Centre.
Other useful indicators	<ul style="list-style-type: none"> Risk and continuity indicators are important in ANC. Venereal Disease Research Laboratory (VDRL) test for syphilis and HIV testing coverage shows quality of care. This should be done in the first ANC visit. Haemoglobin testing and anaemia management rates. ANC referrals shows risk detection (and transport availability). Percentage of women getting third ANC shows, continuity of care, which is often related to perceived quality.
Common problems	<ul style="list-style-type: none"> Attendance for pregnancy test or simple registration without history and examination do NOT constitute antenatal care. Women who have started ANC elsewhere, but who come to your facility for follow-up should be counted as follow-up ANC and not first ANC.

Actions to consider	Low coverage means either the strategy for providing ANC needs to be reviewed, to increase access, or the community should be approached to increase awareness through ASHA, VHSC, BCC and so on.
Indicator MH 4: ANC Anaemic and Hypertension testing and management rates A. % ANC moderately anaemic B. % ANC severely anaemic C. % ANC severely anaemic treated rate D. % ANC hypertension new case detection rate E. Eclampsia cases management rate	
Definition	<ul style="list-style-type: none"> Percentage of pregnant women tested to be moderately anaemic (Hb level <11 g). Percentage of severely anaemic pregnant women treated (Hb level <7 g). Percentage of pregnant women tested with hypertension/high blood pressure (BP >140/90).
Numerator	<ul style="list-style-type: none"> A. Pregnant women tested anaemic <11 g. B. Severely anaemic pregnant women treated (Hb <7 g). C. Pregnant women detected (BP >140/90). D. Number of eclampsia cases managed during delivery.
Denominator	<ul style="list-style-type: none"> A, B & C = Total ANC registration D = Total deliveries (home + institution)
Rationale	<ul style="list-style-type: none"> E. Testing for anaemia and hypertension is an indicator of quality of ANC services and also detection of important risks associated with preventable mortality. F. Hb <7 g and BP >140/90 is a danger sign for pregnant women and should be managed by, arranging for referral transport and informing the medical officer in-charge in advance.
Data source	ANC/Pregnancy registers maintained by health workers.
Other useful indicators	<ul style="list-style-type: none"> G. ANC hypertension management rate. H. LBW rate is common consequence of anaemia and hypertension. I. Still-birth rate/ PNM affected by anaemia. J. Maternal death due to excessive bleeding is more likely in an anaemic. K. Laboratory equipment availability rate.
Suggested level of use	Sub Centre, Primous Health Centre, Community Health Centre.
Common problems	<ul style="list-style-type: none"> L. BP is often not taken and Hb testing is not done. M. Health Sub Centres do not have BP apparatus and Hb kits. N. Sufficient stock of IFA tablets.
Actions to consider	<ul style="list-style-type: none"> O. Address supply side issues. P. Ensure quality of ANC. Q. Awareness generation among mothers to avail complete and quality ANC services.

Three types of ANC coverage indicators are now described, each having a different data element used as numerator.

ANC First Visit Coverage

Numerator is the data element 'First ANC visit', which is defined as the count of all first visits registered (because a count of first visits gives the total); *Denominator* is total number of pregnant women in the same area (which will be based on demographic estimates).

ANC (First Visit in) First Trimester Coverage

Numerator is the data element 'First ANC visit', within the first 3 months of the pregnancy. This data element is a sub-set of the above 'First ANC visit'. *Denominator* is same as above; total number of pregnant women.

ANC Registered Under Janani Suraksha Yojana (JSY) Coverage

JSY is an incentive scheme of the Government of India for pregnant women, in order for them to make use of the institutional health services. (www.mohfw.nic.in/nrhm/rch/guideline/jsy-guideline)

Numerator is the data element 'First ANC visits', (or new women) registered under JSY. This data element is also a sub-set of the 'First ANC visit'. *Denominator* is same as the above ones, total number of pregnant women in the area.

Within our 'data warehouse' framework, the above three data elements represent data standards, whose values need to be reported from the local level for a required time period/periods, which is typically a month, and captured in the data warehouse. The recommended practice is to register 'atomic' values and data elements, that is to keep the data disaggregated. In this way, indicators and different levels of aggregation may be calculated based on the raw data.

Currently, the typical way to generate these three maternal health data elements is to extract them from the register books, used for recording ANC visits. Here, generally, there will be a column for 'First visits', making it easy to count the number of first visits, by month. As we have seen from the case of Himachal Pradesh, there are three different ways in which these data can be reported and included in the data warehouse:

- Data reported from the local unit and the source of the data on paper forms are captured directly into the 'DHIS2', at block level.
- Data reported on mobile phone using the SMS standard are automatically and 'seamlessly' imported into the data warehouse.
- The ANC registry and visits can be registered in an electronic register, using the DHIS Tracker system. From here the values of the three data elements can be aggregated from the ANC register, by the reporting health facility, and by a time period, to be included in a file of a particular format understood by DHIS, and sent (exported) to it where the data elements and values are imported.

In some countries, there will conventionally be a number of paper forms for data reporting, that will be 'owned' by different health programmes, which more or less overlap, with inconsistencies between them. The process of data standardisation is concerned with harmonising these reporting forms and requirements. The advisable approach to data standardisation is to focus on the needed indicators and on identifying the data elements required to calculate these indicators, as we have seen from the example of India. Ideally, the standardisation should result in a data and indicator dictionary, in line with the Indian example. Before it makes any sense to try

to export and import, or transfer data electronically between systems, such as the electronic ANC register and the data warehouse, the data standards and the definition of the data to be transferred, needs to be sorted out. This is a rather complex task, and in the next chapter, we explore further strategies for data standardisation. Here, we provide a snapshot view of technical standards, for data exchange.

2.3.2 Technical Standards for Data Exchange

Having agreed that standards are required for data to be exchanged between systems, for example, which data 'exactly' to send from the electronic ANC register to the 'data warehouse', the syntactic or technical formats then need to be agreed upon, and standardised. These formats are formally 'rigid' technical protocols, specifying exactly the form in which to transfer data, so that the meaning is not lost. Remember here, that it is not only the value and data element that needs to be transferred, but also the location, which is reporting health facility and time period, need to be specified. The unique identifier for all data element values in the 'data warehouse' is place, that is reporting health facility, time period, which can either be a particular month or week, and 'meaning', that is the data element.

A minimum of two applications involved, need to agree upon shared technical protocols of how the data and format could be written and described, from the side of the sender, so that the receiver of the data can comprehend it. The easiest and typical adhoc way to solve the data inter-change between two or more applications is for the systems wanting to inter-change data, to agree on a custom protocol serving, as a gateway between their systems only. The problem with this approach, however, is that it is very rigid and is difficult for new applications to be included, as they would need to adapt. And may be they would have their own way of doing things that is preferable to them. In a world where everybody uses their own standards; communication will be difficult. The agreed 'best practice' way to approach interoperability, is to apply standards shared by all involved systems. In the area of health information, however, there still are not any easy-to-use shared standards available. Fortunately, this is a situation that may change, as a project led by WHO, is currently developing a general standard for the interoperability between systems inter-changing health data called SDMX-HD (Health Data), which represents a general standard for data transfer. Although it is too early to know if SDMX-HD will become the agreed universal standard for exchange of statistical health data, we believe it is useful to provide a brief overview of this standard in the Box 2.4.

Box 2.4

An overview of SDMX-HD¹

SDMX-HD – Standard Development

Background

The HMN Technical Framework provides a high-level architectural vision of how national HIS development might be implemented, and presupposes that such a HIS would consist of interoperating sub-systems, providing data to a national data repository or warehouse, and stops short of explicitly prescribing a data integration strategy. Three possibilities could be considered:

¹ Contributed by Bob Joliffe, one of the authors of the SDMX.HD standard and a core developer of DHIS2.

1. Homogeneous systems – If all of the services could be provided by a single software package, or by a suite of packages from the same vendor, then interoperability becomes an internal arrangement.
2. Gateways – It could have been possible to build gateways to perform translation and data-loading between the reporting sub-systems and the data warehouse system (for example, DHIS2).
3. Communication using an open data standard.

Each of these approaches present difficulties. The use of homogeneous systems in this context is rarely neither desirable nor possible. Besides the significant risk of creating vendor lock-in situations, the range of health service sub-domains is simply too varied to imagine a single product managing aggregate data, patient level data, HR data plus whatever future requirements that might materialise. Building gateways is certainly possible with one or two systems, but becomes more of a burden as the move to more comprehensive health system coverage exposes the need to build new gateways for each new system. The use of an open standard which can be mandated in procurement guidelines provides the best value proposition for the health system owner, in terms of protection against vendor lock-in and providing a 'plug-and-play' growth path in the future. The benefit of using open XML-based standards for interoperability between disparate systems, is well understood, both in terms of economy and efficiency, as well as broader political and developmental considerations related to state procurement of ICT systems in developing countries. We use the term 'open standard' to indicate standards which are not encumbered with intellectual property constraints and which enjoy legitimacy through open, fair and participatory development and maintenance arrangements by a not-for profit agency. The problem was that, up until 2009/2010, there had been no agreed upon standard for exchange of aggregate health data between systems.

In 2010, the WHO published a standard for exchange of data and metadata, for aggregate health information. The new standard, SDMX-HD,² is an implementation of the ISO SDMX standard (TC154, 2005). SDMX is a mature XML-based standard which has found use in a number of domains for international reporting of statistical data. SDMX-HD is a health domain specific dialect of SDMX, which is specialised to represent health indicators. It is important to understand that before two applications can meaningfully exchange a data item, they must (i) have established a common understanding of metadata, and (ii) use a data format which is commonly understood by both parties. SDMX-HD provides formal rules to facilitate this exchange. This is best understood by a simple example. Note that the example simplifies some of the XML for the purpose of instruction.

Example

An electronic medical record system might collect data on immunisations, which take place during clinical encounters at a facility. The facility might be required to report a monthly aggregate data item called 'Total immunisations', broken down by age group (for example, under 1 year old and over 1 year old), and gender (male and female), to the District HIS.

For this exchange to take place, the HIS needs to define and communicate the metadata for the communication to take place. It does this in SDMX-HD, by producing

² Further information about the SDMX-HD standard is available at <http://www.sdmx-hd.org>.

what is called a Data Structure Definition (DSD) message. The DSD is an xml file or a package of xml files in more complex cases, which defines:

Codelists for items such as data elements and disaggregations. In this simple case, 3 codelists would be required for data elements, gender and age, as well as, an additional codelist for health facilities. An example of the gender codelist might look like:

```
<CodeList id="CL_CATEGORY_SEX" agencyID="SL-MOHS"
version="1.0" isFinal="false">
<Name xml:lang="en">Sex</Name>
<Code value="1">
<Description xml:lang="en">Female</Description>
</Code>
<Code value="2">
<Description xml:lang="en">Male</Description>
</Code>
</CodeList>
```

The Codelist is a generic structure used to represent any coded item in SDMX-HD. The other codelists are not shown here.

KeyFamilies which indicate the way in which data values are grouped and disaggregated. So for example, data items like the example above would be represented in a KeyFamily, which indicates the need for age and gender disaggregation. A simplified KeyFamily for this might look like:

```
<KeyFamily id="KF_419035" agencyID="SL-MOHS" version="1.0"
isFinal="false">
<Name xml:lang="en">KeyFamily for categorycombo sex+ age
group</Name>
<Description xml:lang="en"/>
<Components>
<Dimension crossSectionalAttachGroup="true" conceptRef="FREQ"
conceptSchemeRef="CS_COMMON" conceptVersion="1.0"
conceptSchemeAgency="SDMX-HD" isFrequencyDimension="true" >
</Dimension>
<Dimension crossSectionalAttachObservation="true" conceptRef=
"DATAELEMENT" conceptSchemeRef="CS_COMMON" conceptVersion="1.0"
conceptSchemeAgency="SDMX-HD" codelist="CL_DATAELEMENTS"
codelistVersion="1.0" codelistAgency="SL-MOHS"/>
<Dimension crossSectionalAttachObservation="true" crossSectiona
lAttachGroup="true" conceptRef="FACILITY" conceptSchemeRef="CS_
DIMENSION" conceptVersion="1.0" conceptSchemeAgency="SL-
MOHS" codelist="CL_FACILITY" codelistVersion="1.0"
codelistAgency="SL-MOHS"/>
<Dimension crossSectionalAttachObservation="tru
e" conceptRef="SEX" conceptSchemeRef="CS_DIMENSIONS"
conceptVersion="1.0" conceptSchemeAgency="SDMX-
HD" codelist="CL_CATEGORY_SEX" codelistVersion="1.0"
codelistAgency="SL-MOHS"/>
<Dimension crossSectionalAttachObservation="tru
e" conceptRef="AGE" conceptSchemeRef="CS_DIMENSIONS"
```



```

conceptVersion="1.0" conceptSchemeAgency="SDMX-
HD" codelist="CL_CATEGORY_AGE" codelistVersion="1.0"
codelistAgency="SL-MOHS"/>

<TimeDimension crossSectionalAttachGroup="true
" conceptRef="TIME_PERIOD" conceptSchemeRef="CS_COMMON"
conceptVersion="1.0" conceptSchemeAgency="SDMX-HD">

<TextFormat textType="ObservationalTimePeriod"/>

</TimeDimension>

<PrimaryMeasure conceptRef="OBS_VALUE" conceptSchemeRef="CS_
COMMON" conceptVersion="1.0" conceptSchemeAgency="SDMX-HD">

<TextFormat textType="Decimal"/>

</PrimaryMeasure>

</Components>

</KeyFamily>

```

The XML formatting is quite verbose and not easily consumable by a human reader. But the strict conventions used, are quite easy for a programme (such as the electronic medical record system in the example), to use this set of metadata definitions to codify and produce data element values, which can be exported to the district HIS. The metadata defined in the DSD message, would typically represent the reporting requirements within a national HIS. Human management users would not ever interact with the data at this level, and would define the data elements and codelists within the more comfortable context of a user friendly HIS. The DSD is generated from that system.

The data values for the elements to be reported by the electronic medical record system are formatted using a separate type of data message. There are a number of variants on this format, mostly XML, also including CSV. A typical snippet from a data message for which uses the SDMX-HD cross-sectional data format might look like:

```

<DataSet reportingBeginDate="2010-02-01"
reportingEndDate="2010-02-28"

datasetID="OMRS_export" dataProviderID="34">

<Group TIME_PERIOD="2010-02" FREQ="M">

<Section>

<OBS_VALUE AGE="1" SEX="1" value="20" FACILITY="21"
DATAELEMENT="8"/>

<OBS_VALUE AGE="1" SEX="0" value="13" FACILITY="21"
DATAELEMENT="8"/>

.. etc

</Section>

</Group>

<DataSet>

```

The coded values used, are the values from the codelists which were exchanged in the DSD.

Conclusion and Onwards

SDMX-HD represents a distillation of compromise and best practices, both from the information science as well as the HIS domains. It continues to be developed on an iterative basis, with the experience of implementation being fed back into the standard development process. The formal aspect of structuring metadata and data messages, has matured to a point that it has been successfully implemented in a number of independent health information software packages. The longer term goal of the WHO and the national implementation efforts are, to rationalise common codelists to be used across and within countries, and to build tools for better governance of such metadata. Metadata governance, refers to the human processes around publication and revision of code lists, key families, facility lists, that ensure the functionality and flexibility of the HIS over time.

The first concrete application of this standard has been to ensure the interoperability between the following three applications:

- *DHIS2*: Database or data warehouse application of aggregate data.
- *OpenMRS*: Medical records database.
- *iHRIS*: Human resource management system containing records on all the employees.

The interoperability in this case, is to aggregate data in OpenMRS and iHRIS, and to export this to the DHIS2. Though, it would have been easier to implement custom protocols between DHIS2 and OpenMRS and between DHIS2 and iHRIS; the decision was early on, to make use of the emerging standard, SDMX-HD, for this purpose. While the initial implementation is more challenging, using the emerging standard, we ensure its ease, to link with other similar standard-compliant systems in the future. The SDMX-HD is ideally planned to be a general standard used for data transfers and interoperability of aggregate data between software applications, in the health domain. In the case of the integrated 'data warehouse' framework and IHIA, the SDMX-HD will serve as a 'plug-in' functionality, enabling an evolutionary development, where new systems and modules are plugged into the 'data warehouse' using the SDMX-HD standard.

Here we have used the SDMX-HD to illustrate the role and importance of standards for data exchange when developing information systems within an integrated architecture framework where interoperability is a key building block. In reality, however, developing new standards are complicated processes and at this point, it is far from sure that the SDMX-HD exercise will develop into a widely accepted standard.

Summary

Key concepts that can be taken from this chapter are summarised below:

1. The framework for IHIA, includes three levels:

- The social system level
- The application level
- The data level

These levels represent our approach to operationalising the vision of a IHIA conceptualised in Chapter 1.

2. Each level of the HIA, draws upon the services provided by the level(s) below.
3. The social system level, is defining to the whole IHIA, given our focus on the information needs for decision; being the 'raison d'être' or reason for existence of a IHIA.
4. Users' needs, integrated and not fragmented information (as is typically the case), vary with levels and purpose. Information when collected must have a purpose, which is to support action.
5. The application level of the IHIA is best approached through the 'data warehouse', which represents a data repository that can manage data from multiple sources and application domains, as illustrated in the case of Himachal Pradesh.
6. The SDMX.HD represents a current global effort towards developing interoperability data standards.

Reference

Vernadat, F (2007). Interoperable Enterprise Systems: Principles, Concepts and Methods, *Annual Reviews in Control*, 31, 1, 137-145.